E C O N O M I C S   B U L L E T I N

Local Economic Structure and Industry Development in Germany, 1993−2001

Jens Suedekum
University of Konstanz

Uwe Blien
Institute for Employment Research (IAB)

Abstract

This paper formulates a weighted regression approach to analyze the impact of dynamic MAR− and Jacobs−externalities on local employment growth in Germany between 1993 and 2001. We find that Jacobs−externalities matter both in manufacturing, and service industries. MAR−externalities are present only in services.

The authors thank Friedrich Breyer, Pierre−Philippe Combes, Wolfgang Franz, Johannes Ludsteck, Joachim Moeller, Winfried Pohlmeier and Katja Wolf for several very useful suggestions and discussions. All errors and shortcomings are our full responsibility, however.
Citation: Suedekum, Jens and Uwe Blien, (2005) "Local Economic Structure and Industry Development in Germany, 1993−2001." *Economics Bulletin,* Vol. 15, No. 17 pp. 1−8
Submitted: June 7, 2005. **Accepted:** June 24, 2005.
URL: http://www.economicsbulletin.com/2005/volume15/EB−05O40009A.pdf

1. Introduction

The seminal papers of Glaeser et al. (1992) and Henderson et al. (1995) have launched a debate about what economic structure is conducive for the employment growth performance of different industries at the local level. Whereas the former study finds that an industry thrives if it faces a diversified surrounding economic structure, which is consistent with inter-sectoral "Jacobs-externalities", the latter argues that externalities are mainly of the intra-sectoral "Marshall-Arrow-Romer (MAR)"-type.

Although this literature has considerably grown in the aftermath (see Combes/Overman, 2004 for a survey), the present paper is the first study on Germany to our knowledge. In order to facilitate comparisons, we build on the influential study by Combes (2000) on local employment growth in the neighbouring country, France. A close replication of this methodology on German data leads to quite similar results, namely that there is hardly any evidence for dynamic externalities. However, we show that this estimation approach suffers from inherent heteroskedasticity, and is not appropriate in a regression where the response variable is a growth rate. To tackle the problem we develop a weighted regression approach, which is the methodological contribution of the present paper.

When taking this modified model to the data, qualitative conclusions change compared to the unweighted regression. We find that Jacobs-externalities matter in manufacturing and service industries, whereas MAR-externalities are present only in services. The slim evidence on externalities might therefore simply be an artefact from an inconsistent estimation approach.

The rest of the paper is organized as follows. In section 2 we introduce our data set, the specification of variables, and the estimation approach. In section 3 we present the results. Section 4 concludes.

2. The empirical model

2.1. Data

We use the official employment data provided by the German Federal Employment Agency. This information is highly reliable and covers the complete population of all full-time employment relationships (subject to social security) in two years, 1993 and 2001. Employment is observed in 438 NUTS3-districts and in 15 different manufacturing and 10 service industries[1]. For every local industry we know the total employment level and the employment shares in small (<20 workers), medium-sized (20-99) and large (>100) establishments. Furthermore, we know the number of active firms in every cell ($firms_{z,s}$) and the size of every district in square kilometres ($area_z$).

To our knowledge this is the most comprehensive, recent and accurate data set that has been used in this type of study. Since our observation period is still relatively short, we do not test the *timing* of externalities. This issue is taken up in Henderson (1997) and Combes/Magnac/Robin (2004) by means of panel estimation. In this paper we take the common approach and compute a cross-section of growth rates (between 1993 and 2001), and regress these on base year variables that reflect local economic conditions. In this respect, the externalities are

[1] More details about the data set (lists of the districts and industries, summary statistics etc.) are available upon request from the authors.

thought of as being *´dynamic´* rather than *´static´*. Moreover, we are restricted to use employment information and can not construct productivity data. Dekle (2002) and Cingano/Schivardi (2004) have argued that this might influence results, as externalities might be relevant for TFP and output growth, although they do not influence employment growth.

2.2. Specification of variables

For the sake of comparability of results, we closely follow the variable specification of Combes (2000). The dependent variable is the long-run employment growth rate of sector s in district z, relative to the national growth rate of the same industry,

$$y_{z,s} = \frac{(emp_{z,s,2001} / emp_{z,s,1993})}{(emp_{s,2001} / emp_{s,1993})} \tag{1}$$

The exogenous variables are all computed for the base year. Dynamic MAR externalities are identified by the local relative to the national employment share of sector s,

$$spe_{z,s} = \left[\left(emp_{z,s} / emp_z \right) / \left(emp_s / emp \right) \right]. \tag{2}$$

Diversity is measured by a modified (relative) Herfindahl-Hirshman-index that increases with local diversity faced by sector s.

$$div_{z,s} = \frac{1 \Big/ \sum_{s'=1, s' \neq s}^{S} \left(emp_{z,s'} / (emp_z - emp_{z,s}) \right)^2}{1 \Big/ \sum_{s'=1, s' \neq s}^{S} \left(emp_{s'} / (emp - emp_s) \right)^2} \tag{3}$$

It reaches a maximum when all surrounding industries account for an identical employment share. A positive coefficient associated with $div_{z,s}$ signals Jacobs-externalities.

As additional exogenous variables we enter the employment density in region z, $den_z = emp_z / area_z$ to control for general agglomeration economies, and the (relative) average firm size, $fsize_{z,s} = \left(emp_{z,s} / firms_{z,s} \right) / \left(emp_s / firms_s \right)$. Firm size structure has often been included in this type of study to analyze the growth impact of local competition (see Glaeser et al., 1992). Combes (2000) has argued that this identification is problematic, as $fsize_{z,s}$ rather measures the impact of internal economies of scale. To capture competition he includes a dispersion index of firm sizes in every local industry. This is the only variable that we can not construct, since we can not observe plant level employment. As an alternative we use the (relative) employment share in small firms

$$small_{z,s} = \frac{\left(emp_{z,s}^{SMALL} / emp_{z,s} \right)}{\left(emp_{s}^{SMALL} / emp_s \right)}. \tag{4}$$

This variable should reflect local product market competition in the sense that competition is stiffer the higher is the employment share in small firms. Note that all local variables are normalized, so that industry developments at the national level are taken into account.

2.3. Estimation approach

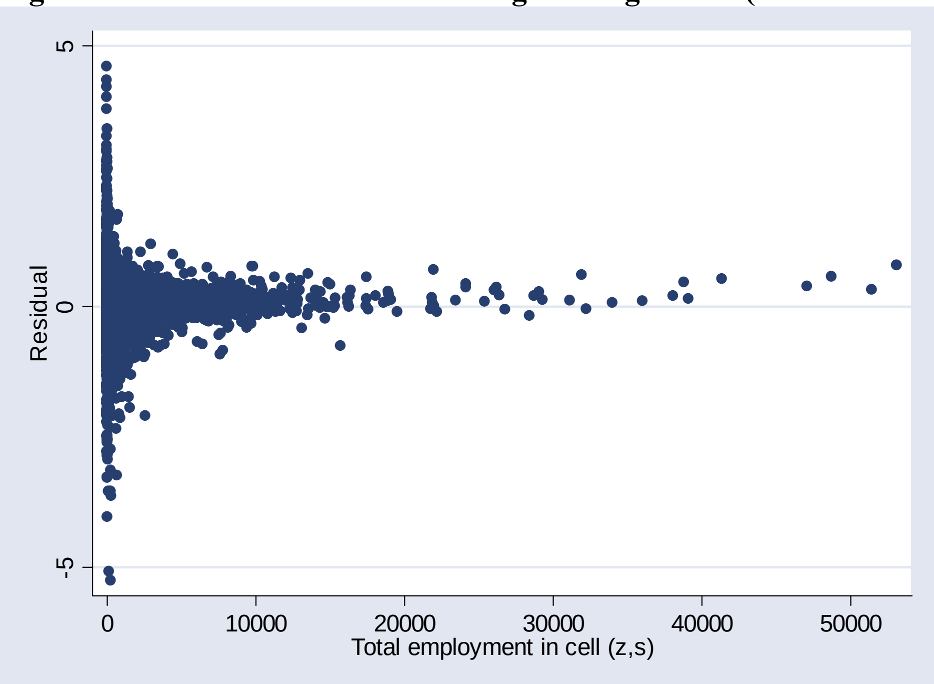
The standard approach in this type of analysis is a simple OLS regression of the form[2]

$$\log(y_{z,s}) = I + \alpha_1 \cdot \log(spe_{z,s}) + \alpha_2 \cdot \log(div_{z,s}) + \alpha_3 \cdot \log(den_z) + \alpha_4 \cdot \log(fsize_{z,s}) + \alpha_5 \cdot \log(small_{z,s}) + \varepsilon_{z,s} \quad (5)$$

We have estimated (5) with robust standard errors separately for manufacturing and services. However, this unweighted estimation (5) suffers from inherent heteroskedasticity due to the so-called "*shipbuilding in the midlands*"-problem.

The issue is the following: The largest district-industry in our data set with 11,779 observations has a total size of 105,675 employees in 1993 (commerce in Hamburg). On the other hand, there is a bunch of very small district-industries (1,238 with less than 100, and 369 with less than 20 employees). Small changes in absolute employment can imply exorbitant jumps in the growth rates (the dependent variable) for these mini sectors, and the error term of the estimation will not be spherical. This is neatly illustrated in figure 1 which depicts the residuals of the unweighted regression for the manufacturing sectors, plotted against the total size of the observation.

Figure 1: Residual structure of unweighted regression (manufacturing)



Any standard test for heteroskedasticity strongly rejects the hypothesis of a constant variance of the residuals. There are various ways to address this issue. Only estimating with robust standard errors and relying on the presumption that the coefficients will be unbiased is not

---

[2] Combes (2000) does not use OLS, since his data suffers from a censoring problem (the truncation of plants with fewer than 20 employees). He employs a generalized Tobit method and then uses ML-estimation in the second stage. Since censoring is not an issue for us, the replication of his unweighted regression approach is done by OLS.

enough in the present context. Since we can track the source of heteroskedasticity, it is preferable to adopt an appropriately specified generalized (weighted) least squares procedure.

Our weighting scheme is derived from a different argument why the unweighted approach is flawed. Recall that the dependent variable of the reduced form equation (5) is (approximately) a growth rate. The (log-)linear specification of the model implies that the growth rate of the aggregate variable $\tilde{Z}$ (aggregate employment growth in Germany) can be built by the arithmetic mean of the growth rates of the single sub-units (the local industries), $\tilde{Z} = \left(1/N\right) \cdot \sum_{j=1}^{N} \tilde{z}_j$. This is in general incorrect, however, which can be demonstrated easily. Let Z_t be the value of an aggregate variable at time t that consists of two components, $Z_t=x_t+y_t$. The growth rate $\tilde{Z}_{t+1} = (Z_{t+1} / Z_t) - 1$ is not equal to the arithmetic mean of the growth rates of x_t and y_t. It is rather given by the *weighted sum* of the growth rates of the sub-units $\tilde{Z}_{t+1} = g_X \cdot \tilde{x}_{t+1} + g_Y \cdot \tilde{y}_{t+1}$, where $g_x = x_t/Z_t$ and $g_y = y_t/Z_t$.

Stated differently, in an unweighted regression approach the first normal equation of regression analysis is violated, according to which the regression hyperplane passes through the point of means of the data (Greene, 1997:238f.). In the present context this additivity property is particularly important, since we consistently divide total employment in Germany into disjunctive subunits and compute several variables by referring to the district level, which itself is obtained by aggregating all s industries in area z. For example, we use relative employment shares that require information about the size of local industries, entire districts and entire industries at the national level. Although our observation units are local industries, conclusions about the impact of specialization and diversity on growth are only possible by aggregating up to the regional and the national level. Consistency then requires that the additivity of growth rates is satisfied.[3]

We will therefore weight the entire estimation equation with a factor $g_{z,s}$ that is given by the employment level of every local industry divided by aggregate employment in all manufacturing (service) industries (emp_{aggr}). This type of weighting procedure, after which the first normal equation of regression analysis is again satisfied, has initially been proposed by Buck/Atkins (1976) in a similar context, and was later extended by Patterson (1991). The weighting is also appropriate in dealing with the heteroskedasticity problem by attaching each district-industry with a weight that reflects the respective importance for aggregate employment. In sum, we estimate the following model

$$\begin{aligned} g_{z,s} \cdot \log(y_{z,s}) = g_{z,s} \cdot I + \alpha_1 \cdot g_{z,s} \cdot \log(spe_{z,s}) + \alpha_2 \cdot g_{z,s} \cdot \log(div_{z,s}) \\ + \alpha_3 \cdot g_{z,s} \cdot \log(den_z) + \alpha_4 \cdot g_{z,s} \cdot \log(small_{z,s}) \\ + \alpha_5 \cdot g_{z,s} \cdot \log(size_{z,s}) + \tilde{\varepsilon}_{z,s} \end{aligned} \qquad (6)$$

where $g_{z,s} = emp_{z,s} / emp_{aggr}$, $\tilde{\varepsilon}_{z,s} = g_{z,s} \cdot \varepsilon_{z,s}$, and $\mathrm{cov}\left(\tilde{\varepsilon}\right) = \mathbf{\Omega}$.[4] This econometric approach (6) is equivalent to a standard GLS-procedure (Greene, 1997:507ff.).

[3] This is a difference with the growth regressions á la Barro/Sala-i-Martin (1995) that, figuratively speaking, also attach the same weight to the United States and Luxemburg. However, in these regressions only country-specific explanatory variables are used, but no control variables that would require aggregation across countries.

[4] Alternatively we can define the matrix W as the diagonal matrix of the weights $g_{z,s}$. The variance/covariance-matrix of the error term $\varepsilon_{z,s}$ from equation (5) is then given by $\mathrm{cov}(\varepsilon)=W\Omega W$. One has to keep in mind that with a weighted intercept the R^2 is no longer defined in the range between zero and one.

3. Results

As a reference, we present the results of Combes (2000) for France in columns 1 and 2 of table 1. He finds that Jacobs-externalities are present in service industries. In manufacturing, diversity even reduces growth. With respect to MAR-externalities he finds that local overrepresentation of an industry significantly reduces employment growth. Considering the results of the unweighted regression (5) for Germany that are presented in columns 3 and 4, we get a quite consistent picture. There is counter-evidence on MAR-externalities. For Jacobs-externalities we find no evidence, neither for manufacturing nor for services industries. These results, based on an unweighted and heteroskedastic regression, would thus lead to the conclusion that the local economic structure hardly matters for the employment growth performance of the different industries.

Focussing now on the results in columns 5 and 6, which are based on the weighted regression (6), we now find clear evidence for the importance of the local economic structure. For manufacturing industries, industrial diversity and thus Jacobs-externalities matter significantly. Yet, the result remains that local overrepresentation reduces growth. The change in conclusions is even more drastic for service industries. We now find evidence both for dynamic Jacobs- and MAR-externalities. Local overrepresentation in 1993 led to significantly faster growth of service industries. On top of that, service industries also benefited from local diversity.

The conclusions with respect to density and the firm size structure are robust with respect to estimation approach and are in line with the results for France. In particular, density has a significantly negative effect on employment growth. Combes (2000) takes this finding as evidence for congestion in dense places. We subscribe to this interpretation, which is consistent with the observation that many countries experience a general suburbanization and de-glomeration process where, in particular, manufacturing employment secularly shifts away from dense city centres to surrounding areas. Average firm size and the variable small that proxies the impact of local product market competition negatively affect growth.

4. Conclusion

Dynamic externalities play an important role for local employment growth in Germany. Using a weighted regression approach we find that manufacturing sectors grow more rapidly if they face a relatively diversified industrial environment, which is consistent with Jacobs-externalities. For service sectors we find evidence both for dynamic Jacobs- and MAR-externalities.

From a policy perspective, our findings cast some doubts on a regional development strategy that aims at supporting “regional clusters” in manufacturing. To be successful, such structural policies would require MAR-externalities, as the basic idea is that regional concentration of an industry will lead to a growth takeoff. In this paper, however, we find no evidence that manufacturing industries grow faster if they locally concentrated. The policy of “regional clusters” might only be successful for service industries, where MAR-externalities matter.

References

Barro, R. and X. Sala-i-Martin (1995), *Economic Growth*, Cambridge: MIT Press.

Buck, T. and M. Atkins (1976), The Impact of British Regional Policy on Employment Growth, *Oxford Economic Papers* 28: 118-132.

Cingano, F. and F. Schivardi (2004), Identifying the Sources of Local Productivity Growth, *Journal of the European Economic Association* 2: 720-742.

Combes, P. (2000), Economic Structure and Local Growth: France 1984-1993, *Journal of Urban Economics* 47: 329-355.

Combes, P., T. Magnac and J. Robin (2004), The Dynamics of Local Employment in France, *Journal of Urban Economics* 56: 217-243.

Combes, P. and H. Overman (2004), The Spatial Distribution of Economic Activities in the European Union, in: Henderson, V. and J.-F. Thisse (eds.), *Handbook of Urban and Regional Economics*, Amsterdam: Elsevier-North Holland.

Dekle, R. (2002), Industrial Concentration and Regional Growth: Evidence from the Prefectures, *Review of Economics and Statistics* 84: 310-315.

Glaeser, E., H.Kallal, J. Scheinkman and A. Shleifer (1992), Growth in Cities, *Journal of Political Economy* 100: 1126-1152.

Greene, W. (1997), *Econometric Analysis*, 3rd ed., New Jersey: Prentice Hall.

Henderson, V. (1997), Externalities and Industrial Development, *Journal of Urban Economics* 42: 449-470.

Henderson, V., A. Kuncoro and M. Turner (1995), Industrial Development in Cities, *Journal of Political Economy* 103: 1067- 1090.

Patterson, M. (1991), A Note on the Formulation of a Full-Analogue Regression Model of the Shift-Share Method, *Journal of Regional Science* 31: 211-216.

Table 1: Regression Results

	FRANCE		GERMANY				
			Unweighted regression		**Weighted Regression**		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	
	Manufacturing n=6664	**Services** n=5842	**Manufacturing** n=6399	**Services** n=4380	**Manufacturing** n=6399	**Services** n=4380	
diversity	-0.051	0.058	-0.0068 (0.802)	0.0002 (0.988)	**0.1349** (0.000)	**0.0296** (0.057)	**diversity**
specialisation	-0.088	-0.211	**-0.0345** (0.005)	**-0.0794** (0.000)	**-0.0460** (0.000)	**0.1173** (0.000)	**specialisation**
density	-0.161	-0.040	**-0.0577** (0.000)	0.0065 (0.081)	**-0.0536** (0.000)	**-0.0054** (0.025)	**density**
fsize	-0.154	-0.110	**-0.1522** (0.000)	**-0.1739** (0.000)	-0.0110 (0.141)	**-0.2589** (0.000)	**fsize**
competition	-0.030	-0.011	**-0.0232** (0.000)	**-0.0594** (0.002)	**-0.0397** (0.000)	**-0.0997** (0.000)	**small**
intercept	0.185	-0.018	**0.2262** (0.000)	-0.0349 (0.052)	**0.2868** (0.000)	0.0125 (0.409)	**intercept**
Likelihood	-17502.72	-14576.76	0.1016	0.1262	0.1689	0.1610	R^2
Likelihood only intercept	-18637.68	-15736.09					

Shaded cells: insignificant at 5% level